

11/26 nothing  
done to finish yet

## Roles of Electronic Publishing on Campus

Brewster E. Kahle  
Thinking Machines Corporation  
Sept 20, 1991  
Brewster@think.com

### Abstract:

-----

Campus Wide Information Systems (CWIS) are being designed and implemented at a number of universities to serve students and faculty with online information easily accessible through the various personal computers and workstations available on campus. Most of these systems are menu-based, largely read-only and centrally organized. Wide Area Information Servers (WAIS) is an electronic publishing architecture, based on multimedia files, that supports distributed information servers for remote question-answering. WAIS is being used for some of these campus information systems since it offers many of the same capabilities, while going further in constituting a backbone for decentralized information distribution. This paper will discuss how a WAIS system could be constructed to give students and professors an even more powerful information tool in the university environment.

### Introduction:

-----

Campus wide information servers have recently become common projects on college campuses to provide information on courses, funding, class, leisure, and other topical subjects to students and faculty. Most of these systems have the following features in common:

- \* they are menu-based for hierarchical browsing (try UNC's system by telnetting to: \*\*\* port);
- \* some have limited client-server functions to provide different interfaces on the dominant computer platforms: Mac, PC, and dumb terminals; (try the gopher system file:\*\*\*)
- \* some use existing products such as VTX from DEC, IBM book reader (\*\*\*? find out from Jim Fullton) that are proprietary, centralized, and closed;
- \* they are usually developed and maintained by the computer support staff;
- \* they attempt to integrate the online library catalog.

One feature seems to be common to all of them: since they are actively being developed in the university setting, there is an openness to sharing code and servers amongst universities.

A setting of open access and shared resources can be a fertile environment and a unique opportunity for the development of electronic publishing (publishing over wires), and for establishing and refining an open system of servers and clients.

Wide Area Information Servers (WAIS) is a system for distributed servers and clients to share information in any format. Even though WAIS has not yet developed a menu-based browsing interface on many platforms (the closest is the dumb-terminal interface done by John Curran usable via telnet to quake.think.com with login wais), the open nature of the system and the mature freeware versions of many of the pieces has lead some CWIS implementors to use parts or all of it in their systems (such as Jim Fullton at University of North Carolina, Fullton@rhumba.oit.unc.edu). This paper will discuss how such an electronic publishing system could be used to enhance the current model of campus information systems to provide an innovative, useful tool for students and professors.

### WAIS Overview

-----

The Wide Area Information Servers project was started by Thinking Machines Corporation to construct an electronic publishing system for

non-professional searchers. It is based on a client-server model using a standard protocol (an extension of NISO Z39.50) to allow users to find and retrieve information from a large number of servers. To make the system friendly to end users, natural language has been used as the query format, while the retrieved information can be anything from text and pictures, to video and formatted records. For over five years, different parts of this system have been used in various situations, and have been widely accepted by end-users.

The system is composed of three main parts: a client program provides the user interface; servers do the indexing and retrieval of documents; and a standard protocol is used to transmit the queries and responses. The client and server talk to each other through the protocol but are at the same time isolated. Any client that is capable of translating a users request into the standard protocol can be used within the system. Likewise, any server capable of answering a request encoded in the protocol can be part of the system.

On the client side, queries are formulated as natural language questions. The client application then translates the query into the WAIS protocol, and transmits it over a network to a server. The server receives the transmission, translates the received packet into its own query language, and searches for documents satisfying the query. The list of relevant documents is then encoded in the protocol, and transmitted back to the client. The client decodes the response, and displays the results. The full documents can then be retrieved from the server, if required by the client.

From the user's point of view, a server is a source of information. It can be located anywhere the user has access to: on the user's own workstation, organization network, a wide area network, or a system connected by modem." The user's workstation keeps track, via the WAIS client, of a variety of information about each server. This public information about a server includes how to contact it, a description of the contents, and the cost. In addition, a special feature of the client program allows individual users to store private information about the servers they use. Some users need to budget the money they are willing to spend on information from particular servers, to decide how often and when each server should be contacted, and to assess the relative usefulness of each server. This information helps guide the user's workstation or personal computer in automatically making cost effective decisions in contacting servers.

One of the primary motivations behind the initial development of the WAIS system was to develop a single interface for accessing multiple sources. In fact, with most current retrieval systems, complications develop as soon as one begins dealing with more than one source of information. If, for example, the same question needs to be asked to various sources, the user has to repeat the query process with each database, most likely using a different query language, a different style of interface, and a different billing system.

With WAIS, the user selects a set of sources to query for information, and then formulates a question. When the question is run, the system automatically asks all the servers for the required information with no further direct interaction on the part of the user. The returned documents are sorted and consolidated in a single list of results to be easily manipulated by the user. This way, the user has transparent access to a multitude of local and remote databases.

One of the most far-reaching aspects of the WAIS project is the development of an open, publicly available protocol. Three companies (Thinking Machines Corporation, Apple Computer, and Dow Jones Inc.) have jointly specified a standard protocol for information retrieval. Ideally this protocol should be internationally standardized, yet flexible enough to adapt to new ideas and technologies, functioning over any electronic network, from the highest speed optical connections to phone lines.

The use of an open and versatile protocol fosters hardware independence. This not only provides for a much wider base of users, it allows the system to seamlessly evolve over time as hardware technology progresses. It also provides incentive to produce the best components possible. For example, the protocol provides for the transmission of audio and video as well as text, even though, most workstations presently are unable to handle these

data formats. However, WAIS clients are free to ignore pictures and sound returned in response to questions, and to display and retrieve only text. This inability, does not hinder higher-end platforms from exploiting their greater processing power and network bandwidth.

The WAIS protocol is based on the existing Z39.50 standard [NISO 1988] extended to incorporate many of the needs of a full-text information retrieval system. To allow future flexibility, the standard does not restrict the query language nor the data format of the information to be retrieved. Nonetheless, a query convention has been established for the existing servers and clients. The resulting WAIS Protocol is general enough to be implemented on a variety of communications systems. We are continuing to work with the standards committee to put the needed functionality into future versions of the Z39.50 standard.

The success of a WAIS-like system depends on a critical mass of users and information services. In order to encourage development and use, Thinking Machines Corporation is not only publishing a specification for the protocol, but is also making the source code for a WAIS Protocol implementation freely available. (Where?) While this software is available at no cost, it comes with no support. We hope that it will facilitate others in developing servers and clients.

#### WAIS applications on Campus:

-----  
Intro paragraph listing uses.

#### LIBRARY CATALOGS ARE GOING FROM CARDS TO TERMINALS TO REMOTE ACCESS:

Remote library catalog access will make libraries much easier to use and pave the path of full-text library systems. The physical library will become an archive and reading room rather than the only access point to its holdings.

Expensive conversions from paper cards to online catalog records (how much?\*\*) has been underway in most major libraries for the last 20 years. Companies like OCLC, that have large online catalog holdings, have helped in this process by providing cards at a small cost (\$.70 to \$2 (ref private conversation with Ralph LeVan of OCLC)) and has created a union catalog of many library holdings.

In the early days, online catalogs were often accessed from terminals in the physical library (see picture) and rarely available remotely. Thus the online system was a direct replacement for the physical cards. They offered quicker access and reduced maintenance costs, but no new features. These systems were so popular that people would stand in line to use them rather than use the physical cards.

As remote access became available, it was largely through a menu interface that was specific to that catalog. Today about 150 US catalogs are available for remote access over the internet (see Art St. George's and Billy's database of online catalogs available over the Internet). The University of California's union catalog system, Melvyl, gets fully 30% of its access through the net rather than patrons using the library systems (ref personal conversation with Clifford Lynch, Director of Library Automation, Office of the President, University of California). There are limitations to these systems, as they are all different and incompatible.

A computer-to-computer protocol for remote card catalog access (Z39.50 ref\*\*) promises to make it easy to search multiple card catalogs using the same interface. WAIS is using this protocol. Some companies that make library software systems are finding support of Z39.50 a good marketing position (ref James Michael of DataResearch private conversation).

Most universities have made their catalogues available campuswide through remote access. WAIS access will make it easy to search multiple catalogs at one time and allow the user to download the needed information for references and further work. As libraries collect full-text or abstracts for their holdings, this material can also be effectively searched and retrieved using WAIS.

## MOVIE SCHEDULES CAN BE USED TO AUTOMATICALLY ALERT YOU TO APPROPRIATE MOVIES:

On university campuses, movie schedules have traditionally been distributed through fliers, school newspapers, and hotline numbers. Movie schedules can now be effectively distributed electronically, with WAIS adding further functionality to this distribution. In the old system, the fliers were usually distributed at shows, in the living groups, and in common areas. They would normally include a short abstract and a small photo to advertise the movie. (find out how many schedules are printed for LSC \*\*\*). Hotlines with recordings of the schedules are now commonly available using inexpensive answering machine hardware. More sophisticated hotline (AudioTex) systems exist in some areas, such as 212-777-FILM in New York City, which ask what films the caller is interested in and gives closest locations, and show times.

WAIS can lead to easy access of changing schedules based on title, actors, directors, subject matter, location, or date. Furthermore, WAIS can be used to alert the user when an appropriate movie is playing rather than always requiring the user to search for the movie. This can be achieved by making a listing of movies with abstracts and still shots available as a WAIS server. The server can then be searched based on any piece of the description or combination of information.

(show WAIS station screen with question "Movies by Kubric or any depressing films", similar to "2001 A Space Odyssey", source "Boston Area Movies", resulting documents: "A Clockwork Orange showing March 15 at Kresge theater, MIT" "Repulsion by Polanski showing May 3 at Wellesley College" With still shot from Clockwork orange and abstract.)

The WAIS system paves the way for a more ambitious general video holdings catalog of lectures and movies, and could be used to distribute the videos when the networks reach the necessary speed. (at North Carolina has such a network now, and plans to start experimenting with this soon).

## CLASS SCHEDULES AND CATALOGS CAN INCORPORATE INDEPENDENT STUDENT EVALUATIONS

Electronic distribution of class schedules and catalogs can decrease paper waste, and incorporate student evaluations, creating a more useful class selection tool. Class catalogs are currently distributed in books that often resemble phonebooks, while the smaller class schedule books are printed for one-time use, right before registration. Student evaluation results are sometimes distributed separately, usually by a different body, making them difficult to use in the class selection process.

Computers not only could aid in finding non-conflicting and available classes, but electronic class lists would allow students to integrate divergent class information from many sources. If student organizations made evaluation information available as a WAIS server, a user could select a course from the catalog, see the schedule, and review comments from previous students through a single, easy to use interface. Since the evaluation information is likely to be controversial, this architecture relieves the school from the awkward position of selecting which evaluations to publish with the catalog. Thus the integration would be done by the user from multiple WAIS servers to find official and unofficial information about courses.

A further step could be online registration, which could show class availability and number of spaces left in a course, and would allow juggling of schedules and rooms to accommodate demand rather than forcing students into a static schedule.

## SCHOOL NEWSPAPERS AND CAMPUS ORGANIZATION NEWSLETTERS

Electronic publication of the school newspaper and organization newsletters could lead to more indepth stories and less dependence on advertising, thus making more useful publications for the students. School newspapers and newsletters are costly to produce and are of limited length based on paper and production costs. For example, the MIT "Tech" prints \*\*\* copies for

the \*\*\* MIT students for \*\*\* per issue, while the only costs are in materials since the entire staff is volunteer. Electronic publishing would provide significant reduction in cost, since most of the processing would be done on computers owned by the school.

The resulting electronic paper and newsletters could be easy to browse, could alert the student to events that would be of personal interest, and could provide students with a wider forum to express their views. Attention to presentation and content is just as important in the digital form as in the paper version, and could start by taking the content of the paper version, and making it available on the CWIS. The capability of electronic systems to search back issues can give life to articles long after all the paper copies have been lost. Teaching students to manage and edit such a publication can be just as professionally helpful as the work on hardcopy school papers and newsletters is today.

#### BUS SCHEDULES, CAMPUS PHONE BOOKS AND OTHER LOGISTIC INFORMATION

Other information useful to students can be incorporated into WAIS such as bus schedules, campus maps, and cafeteria locations. University of North Carolina has already made this type of information available through their menu-based system, but WAIS access might be more successful.

Another case is represented by campus phone books that are usually distributed in book form and have been subject to abuse by mailing list vendors. Electronic distribution of this information through a WAIS server can make up-to-date information cheaper and easier to distribute, and help, somewhat, with the privacy problems.

Questions such as "Bruce in East Campus" are easy to answer for a WAIS server. The server does not have to come up with one right answer, rather a ranked list of possible answers. Since people do not always have quite the right specifications, WAIS can supply further suggestions that are not quite exact matches. If the privacy issues can be handled, it would be nice to have further information about the person's department, home address, and hobbies. Pictures can also aid in making the electronic version of the phone book a useful tool for helping people get to know each other.

Since the distribution of the phone book entries would be done through a WAIS server, the responses or documents could be filtered relative to who was requesting the information. Most important, the server could detect blatant attempts to download the whole directory. This is not to say that electronic publishing will inhibit those that want to abuse that information. In fact, electronic publishing generally makes it easier to copy large quantities of data. However, careful implementation of the WAIS system can protect against some of these abuses.

#### BULLETIN BOARD EXTENSION: ARCHIVE LOOKUP

One of the more interesting applications of WAIS is in conjunction with bulletin board systems to allow easy access to archives. Bulletin board systems (such as Usenet) are popular discussion groups on all sorts of topics and are becoming more and more common on campuses. These systems are more sophisticated than most email systems for following conversations (threads) and filtering incoming messages. They are designed for interactive use of timely information. These systems are rarely augmented, however, with an ability to search the archives that store old messages or important messages. WAIS can provide this sort of archiving and searching system for bulletin boards.

Used correctly, bulletin boards can be a useful tool to circulate information within campuses, which benefits the university and its educational goals. WAIS can add functionality to these systems.

#### GATEWAYS TO OTHER SYSTEMS, BOTH COMMERCIAL AND NON-COMMERCIAL:

In some CWIS implementations, there is a menu item for using other systems that are explicitly added by the system administrators. This type of control allows the maintainers to assure quality, but at the cost of

restricting access to other services. WAIS can allow easy access to other systems without any explicit action on behalf of the systems administrators. This is done by allowing users to probe the WAIS directory of servers to find new services.

The number of future servers is possibly very large, and their subject matter heterogeneous. Minitel, the French videotex system, has 13,000 different information services, and Dialog has 600 databases. Finding the appropriate service can be a search in itself. Adding these services one-by-one to a menu-based system on each campus probably will not scale well enough to suit the student populations.

WAIS is organized to have one or more "directory of servers". They can be searched like any other server to help users find the right servers for particular questions or needs. There can be a directory of servers for a particular campus, as well as a more global server such as the one operated at Quake.think.com by Thinking Machines. Thus, using the WAIS protocol within a CWIS automatically gives a gateway to other systems much as using TCP/IP on one's LAN allows access to other internet services.

OTHER POSSIBILITIES: (it's up to you to decide if you want to develop this part, maybe you should add it to previous paragraphs.....)

WAIS can be used to turn in class assignments, which can help graders write comments and return them promptly.

Really??? I'd at least say "WAIS could be used..."

Email between students can help in group projects and help in working together even though they may live apart and work on different schedules. Cooperative skills are growing more important in our post-manufacturing era.

Electronic communications of all sorts will help separate campus's work together, leveraging the skills and good teachers that may not be available in other areas.

Distribution of freeware and site-licensed software can be effectively done through a WAIS server system.

#### FUTURE:

Just as in the late 1800s when there were mostly horses and few cars, and everyone knew the future was in those rubber tire beasts, it is clear that future information systems will be heavily electronic. Developing and stretching the technology will proceed at a rapid pace in environments that have the networking and computer resources. The differences between those institutions that are employing this technology and those that are sitting it out will become more striking as the trend continues.

Universities and colleges will probably exercise a leading role in electronic publishing technology, since their users are young and flexible, their environment thrives on free information exchange, and a large number of them are linked by a common communications backbone, the Internet.

The revolution will be interesting to watch, but even more fun to participate in. The electronic publishing revolution in the universities will have many participants since universities are an innately distributed server system. WAIS will hopefully be a helpful set of tools for the first electronic campaign towards decentralized, user controlled, and dynamic information exchange.

*References*

it is clear that future information systems will be heavily electronic. Developing and stretching the technology will proceed at a rapid pace in environments that have the networking and computer resources. The differences between those organizations that are employing this technology and those that are sitting it out will become more striking as the trend continues.

Universities and colleges will be probably be a leader in electronic publishing technology since the users are young and flexible, there is an environment of free information exchange, and the presence of the Internet as a communications backbone. The revolution will be interesting to watch, but even more fun to participate in. The electronic publishing revolution in the universities will have many participants since it is an innately distributed server system. I hope that WAIS can be a helpful set of tools for the first electronic campaign towards decentralized, uncontrolled, and dynamic information exchange.

#### REFERENCES:

[FULLTON] Jim Fullton (fullton@lambada.oit.unc.edu) at University of North Carolina is using WAIS as the backbone of their CWIS system.

#### ACKNOWLEDGEMENTS:

Thank you to Jim Fullton, Harry Morris, Jonathan Goldman, Barbara Lincoln.

\*\*\*\*\*

Date: Thu, 19 Sep 91 08:41:09 EDT  
From: Jim Fullton <fullton@rhumba.oit.unc.edu>  
To: brewster@Think.COM  
Subject: paper

I'm looking over it now - looks really good.

#### Introduction:

You might make a note about faculty/staff usage of CWIS - with the current implementations, it is frequently much heavier than student use. This is unfortunate, but most systems are rather awkward to use. More flexible systems would change this.

Menu-based browsing in WAIS should be the access mechanism of last resort - for those with dumb terminals or slow dial-up lines. The ideal CWIS will have canned clients for many different systems available. These clients could be installed in a matter of minutes to get the user going. They would also be installed in student computing labs found in dorms and teaching buildings.

#### WAIS Overview

Looks great. You might figure out some way of mentioning "free security" - instead of charging users, the server would authenticate them. Useful for medical environments, etc. We have had a lot of interest from our school of medicine in that.

The part about a critical mass of users bears repeating. I am cajoling information providers on campus to make their stuff available, while trying to get people to write clients - with those two things, the users will come naturally.

#### WAIS on Campus

Put something in there about researcher support - that is one of our main goals - our research information server will be

geared solely to researcher support - for those who have WAIS!

#### Class Schedules

Many campuses are going to telephone registration - that's how we do it. But a WAIS mechanism for looking at available classes as well as # of free spaces would be very useful.

#### School Papers

All campus organizations could have their newsletters put up on a wais server. Not just the newspaper.

#### Phone Books

We could put fake entries in our phone books. If one of those fakes gets a piece of junk mail, somebody gets a lawsuit. Electronic restrictions would be tough to figure out - limits on the number of accesses could be easily defeated by a determined individual.

#### Other Possibilities

A good WAIS system will get people using computers, period. They could then make use of other features, such as e-mail, etc. Get people using WAIS and other skills will follow.

The site-license software note is particularly good. But, it will require authentication.

How about a "Forum" server? E-mail to it on any subject would automatically be indexed and saved for a few days - just like the bulletin boards on the walls of the Student Union.

Or a "rides" server for carpooling? Or a "jobs" server like ours, but with public posting access?

=====

The paper looks good - would you mind if I showed it to Paul and Judd?

Jim